

We claim:

Sub 11
5 steps of:

1. A computerized method of comparing proper names, comprising the steps of:
 - obtaining data representing a first proper name and a second proper name;
 - converting the data representing said first and second names to first and second pronunciation equivalent phonetic alphabet representations, equivalent to at least respective portions of said first and second names;
 - 10 comparing said first and second pronunciation equivalent phonetic alphabet representations to determine a likelihood that said first and second names represent the same entity; and
 - producing a signal indicating said likelihood that said first and second names represent the same entity.
- 15 2. The method of claim 1 wherein said data representing first proper name and said second proper name is obtained as a string of characters.
3. The method of claim 1 wherein said phonetic alphabet representation
20 is an International Phonetic Alphabet representation.

Sub 2

4. The method of claim 1 including the further step of processing at least one of said first and second names to assign to said name one of a set of categories of likely ethnic origin of said name,
25 wherein said comparison of said first and second phonetic alphabet representations is performed according to an algorithm that varies depending on said assigned category of likely ethnic origin.

09275766-032599

Sub 1
D/L 5.

5

6. The method of claim 4 wherein said algorithm ignores certain portions of said pronunciation equivalent phonetic alphabet representations in the comparison, depending on said assigned category of likely ethnic origin.

10
Sub B3 7.

15

converting at least a portion of the different proper names in the database to a pronunciation equivalent phonetic alphabet representation equivalent to at least a portion of said respective proper names and assigning said pronunciation equivalent phonetic alphabet representation as keys for records in said database corresponding to said different proper names;

receiving data representing the input proper name as a string of characters;

20

converting the data representing said input proper name using the same said pronunciation equivalent phonetic alphabet representation, equivalent to at least a portion of said input proper name;

25

comparing said phonetic alphabet representation of said input proper name to said record keys to determine a likelihood that said input proper name represents the entity associated with said record;

eliminating as potential matches those records for which the likelihood that said input proper name represents the entity associated with said record key falls below a predetermined threshold; and

processing the records remaining after said eliminating step as potential matches for said input proper name.

8. The method of claim 7 comprising the further step of conducting a second database search ~~pass~~ after said eliminating step, using an algorithm which requires greater processing time for each record than said step of comparing said phonetic alphabet representation of said input proper name to said record keys, to further evaluate the likelihood that said input proper name represents the entity associated with each record.

9. The method of claim 7 wherein said phonetic alphabet representation is an International Phonetic Alphabet representation.

10. The method of claim 7 including the further step of processing said input proper name to assign to said name one of a set of categories of likely ethnic origin of said input proper name, wherein said comparison of said phonetic alphabet representation to said record keys is performed according to an algorithm that varies depending on said assigned category of likely ethnic origin.

11. The method of claim 10 wherein said algorithm compares different portions of said pronunciation equivalent phonetic alphabet representation of said input proper name to said record keys, depending on said assigned category of likely ethnic origin.

12. The method of claim 10 wherein said algorithm ignores certain portions of said pronunciation equivalent phonetic alphabet representation of said

00276766-032600

Sub 84

Sub ⁵BS

13. A name processing and matching system comprising:
a database containing a plurality of proper name database entries and records associated respectively with said proper names;
database processing means associated with said database for converting at least a portion of each of said plurality of proper name database entries using a pronunciation equivalent phonetic alphabet representation, equivalent to a respective portion of said proper name database entries;
input receiving means for receiving data representing an input proper name;
phonetic processing means associated with said input receiving means for converting at least a portion of said data representing said input proper name using said pronunciation equivalent phonetic alphabet representation;
comparison means associated with said database processing means and said phonetic processing means for comparing said pronunciation equivalent phonetic alphabet representation of said input proper name to said pronunciation equivalent phonetic alphabet representations of said proper name database entries to determine a likelihood that said input proper name represents the same entity as each of said database entries; and
output means associated with said comparison means for eliminating as potential matches those records associated with proper name database entries for which the likelihood that said input proper name represents the entity associated with said database entry falls below a predetermined threshold, and for processing the records remaining after said eliminating function as potential matches for said input proper name.

14. The system of claim 13 wherein said proper name database entries and said data representing said input proper name are strings of characters.

15. The system of claim 13 wherein said phonetic alphabet representation is an International Phonetic Alphabet representation.

16. The system of claim 13 further comprising name classifying means for processing said input proper name to assign one of a set of categories of likely ethnic origin of said input proper name,
10 wherein said comparison means comprises variable processing means for performing said comparison of said pronunciation equivalent phonetic alphabet representation of said input proper name to said pronunciation equivalent phonetic alphabet representations of said proper name database entries according to an algorithm that varies depending on said assigned category of likely ethnic origin.

15 17. The system of claim 16 wherein said variable processing means compares different portions of said pronunciation equivalent phonetic alphabet representation of said input proper name to said pronunciation equivalent phonetic alphabet representations of said proper name database entries, depending on said
20 assigned category of likely ethnic origin.

25 ⁵⁴⁰ 18. The system of claim 16 wherein said variable processing means ignores portions of said pronunciation equivalent phonetic alphabet representation of said input proper name in the comparison, depending on said assigned category of likely ethnic origin.

add A3
Add C27

add D27

add E1

add F2

00275766-032500